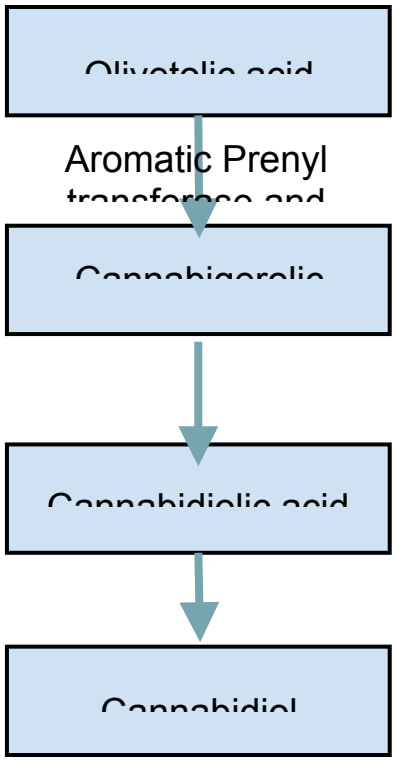
Grace's notebook
3/3/15
Today we had a brainstorming session for project fro this summer. These were the most popular:

- Taq polymerase
- Cannabidiol
- Spider Silk production
- Killing/altering Kudzu
- Teraforming Mars
- Pathogen Detector

3/4/15
I began doing background research on Cannabidiol. First we have to identify the chemical pathway that synthesizes CBD. We need to know the precursors and the enzymes used to determine whether this is a feasible project or not. I found several articles about the uses of CBD oil which indicates that is very useful in the treatment of cancer especially lung and breast cancer. I was also able to find out that cannabidiolic acid(CBDA) is the step directly before CBD, and the CBDA is made from Cannabigerolic acid(CBGA) acted on by Cannabidiolic acid synthase(CBDAS). I haven't found any of the substrates or enzymes before that yet, but this may be far enough back in the pathway for us.

3/6/15
Yousef was able to find the substrate and enzyme that comes before cannabigerolic acid. So we presented all of the information we had gathered up to this point in a power point at our Synthetic Biology Club meeting today. The chemical pathway as best as we known know is as follows:



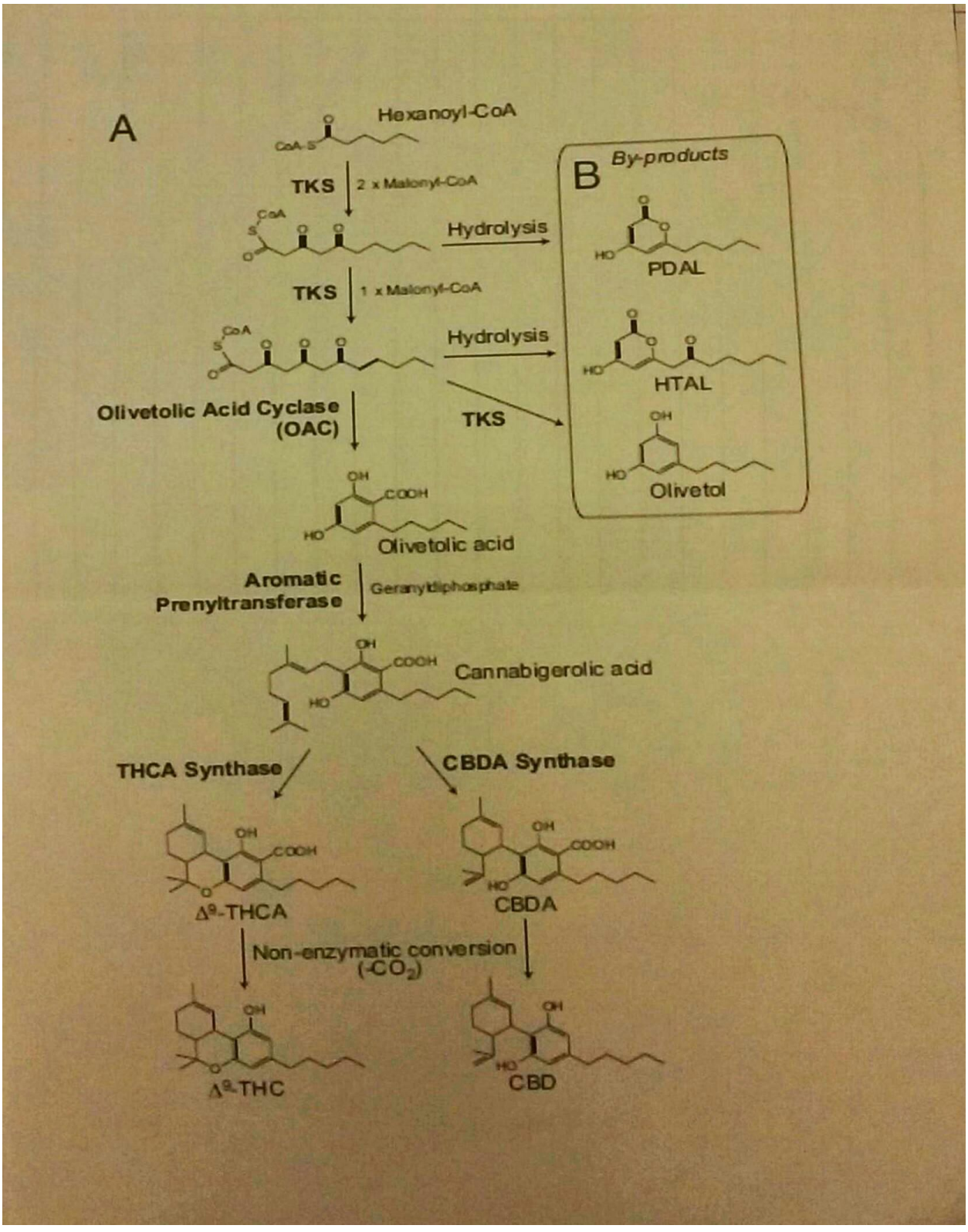
3/30/15

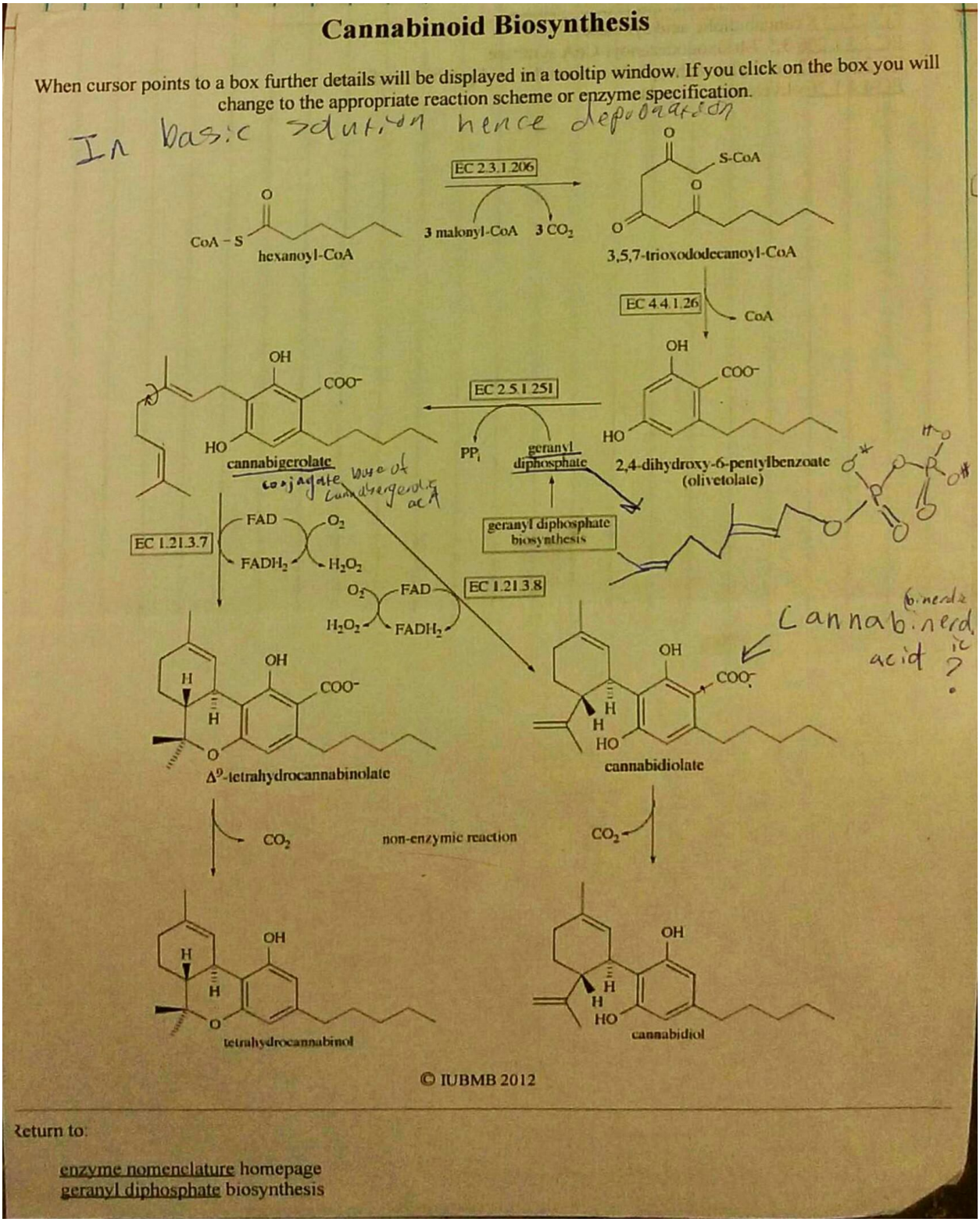
Did more research on Cannabidiol. I found the cDNA sequence for CBDA synthase. Also discovered that Cannabigerolic acid may be purchasable. One of the companies is in Denmark though which may make it impossible for us to order from them.

4/1/15

I met with Dr. Brewer about possibly purchasing CBGA. He said that it isn't purchasable through any of the companies we usually use. Trying to purchase it from a different company may be possible, but would be more difficult.

We also examined the pathway more closely because there appear to be some differences in the chemical structures on the graphic of the chemical pathway and on the websites we have looked at. Dr. Brewer suggested I meet with one of the Organic Chemistry professors to determine if there is in fact a difference. We can also get their advice on where in the chemical pathway would be best to start as well as if they think the project is feasible at all. Derrica suggested I contact Dr, Harden. He agreed to see us on Monday, 4/16/15, at 9:30 in his office. In preparation for the meeting I printed out the two versions of the CBD pathway and have attached them here


A
Hexanoyl-CoA
TKS
2 x Malonyl-CoA
Hydrolysis
TKS
1 x Malonyl-CoA
Hydrolysis
Olivetolic Acid Cyclase (OAC)
TKS
Olivetolic acid
Aromatic Prenyltransferase
Geranyldiphosphate
Cannabigerolic acid
THCA Synthase
CBDA Synthase
Δ9-THCA
CBDA
Non-enzymatic conversion (-CO2)
Δ9-THC
CBD
B
By-products
PDAL
HTAL
Olivetol



4/6/15

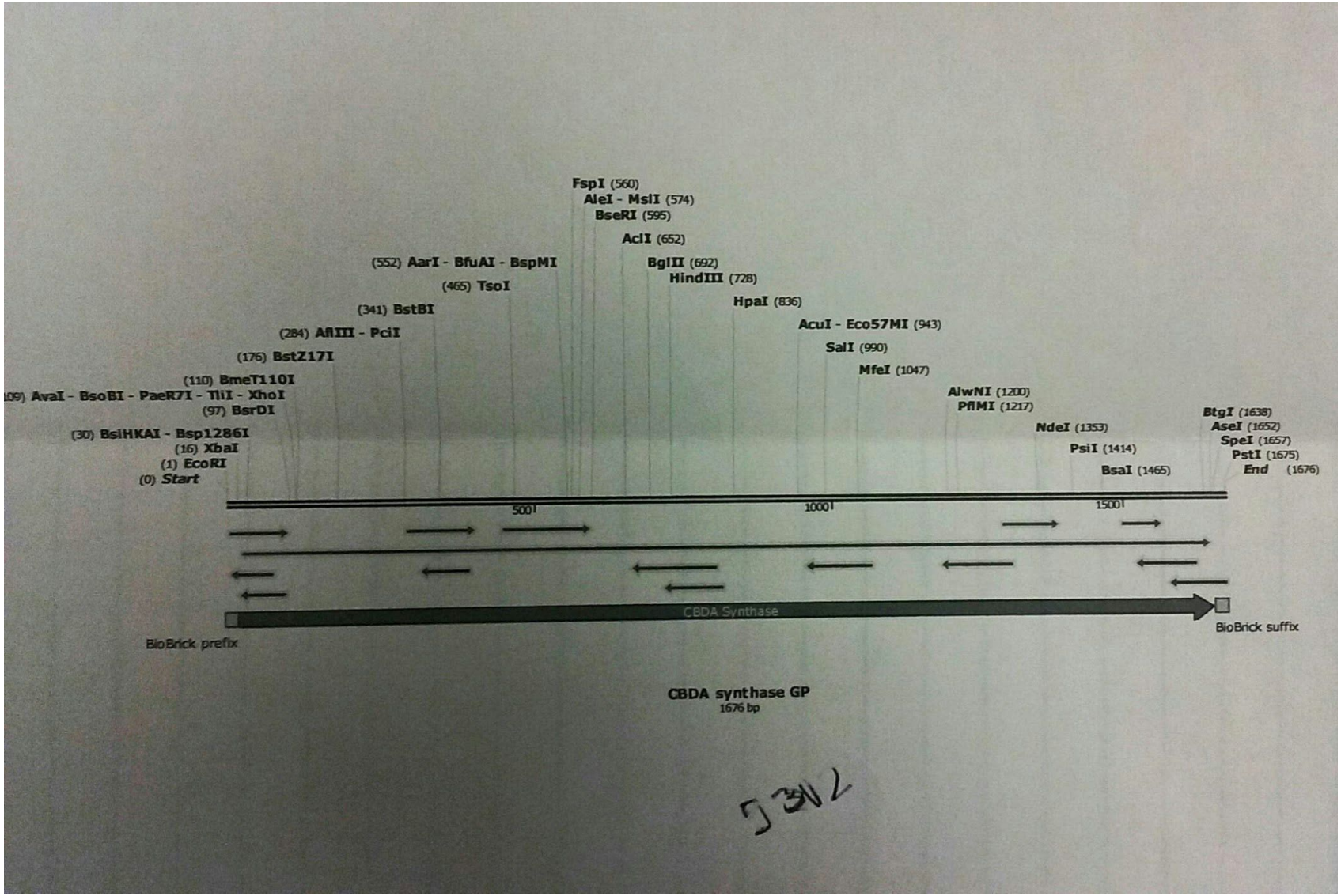
Derrica and I met with Dr. Harden this morning to discuss the differences in the 2 CBD pathways we found and our project in general. He told us that differences in the chemical structures were merely artistic. He also explained that the second pathway we found was in

basic solution so the acids were deprotonated, and that it was more detailed. Dr. Harden also agreed that we should start at the CBGA to CBDA step because it is simplest and it is at the crucial branching point between THC and CBD.He suggested that we determine the conversion rate from CBGA to CBDA to ensure that we get high enough yields, and that we talk to a biochemist.

I then met with Dr. Brewer to inform him what Dr. Harden said. He told us to not worry about the conversion rate for now. He asked if I had found the cDNA sequence for CBDA synthase and instructed me to put it into snap gene with the igem prefix and suffix so that we can send it to IDT.

4/7/15

Derrica and I put the cDNA sequence for CBDA synthase into snap gene with the IGEM prefix and suffix. We saved in on the desktop for easy access. A hard copy is attached below.


FspI (560)
AleI - MslI (574)
BseRI (595)
AclI (652)
(552) AarI - BfuAI - BspMI
BglII (692)
HindIII (728)
(465) TsoI
HpaI (836)
(341) BstBI
AcuI - Eco57MI (943)
(284) AflIII - PciI
SalI (990)
(176) BstZ17I
MfeI (1047)
(110) BmeT110I
AlwNI (1200)
09) AvaI - BsoBI - PaeR7I - TliI - XhoI
PflMI (1217)
(97) BsrDI
BtgI (1638)
NdeI (1353)
AseI (1652)
(30) BslHKAI - Bsp1286I
SpeI (1657)
PsiI (1414)
(16) XbaI
PstI (1675)
BsaI (1465)
(1) EcoRI
End (1676)
(0) Start
500
1000
1500
CBDA Synthase
BioBrick prefix
BioBrick suffix
CBDA synthase GP
1676 bp

J312

4/13/15

Cell Biology Unit Web Server.
Bioinformatics:

align results

Please cite: *Pearson, W.R., Wood, T., Zhang, Z., and Miller, W. (1997) Comparision of DNA sequences with protein sequences, Genomics 46: 24-36*

align Search Help

```
>_ CBDA                                                544 aa vs.
>_ THCA                                                545 aa
scoring matrix: , gap penalties: -12/-2
83.9% identity;         Global alignment score: 3165

               10        20        30        40        50        60
395641 MKCSTFSFWFVCKIIFFFFSFNIQTSIANPRENFLKCFSQYIPNNATNLKLVYTQNNPLY
_      .N..A.............L..H..I.............KH....VA.P......HDQ..

               70        80        90       100       110       120
395641 MSVLNSTIHNLRFTSDTTPKPLVIVTPSHVSHIQGTILCSKKVGLQIRTRSGGHDSEGMS
_      ..I.....Q....I..............NN....A.....................A....

              130       140       150       160       170       180
395641 YISQVPFVIVDLRNMRSIKIDVHSQTAWVEAGATLGEVYYWVNEKNENLSLAAGYCPTVC
_      ........V......H.........................I.........FPG......G

              190       200       210       220       230       240
395641 AGGHFGGGGYGPLMRNYGLAADNIIDAHLVNVHGKVLDRKSMGEDLFWALRGGGAESFGI
_      V....S.....A.....................D................I....G.N...

              250       260       270       280       290
395641 IVAWKIRLVAVP-KSTMFSVKKIMEIHELVKLVNKWQNIAYKYDKDLLLMTHFITRNITD
_      .A....K.....S...I.....N....G....F..............V.......K....

     300       310       320       330       340       350
395641 NQGKNKTAIHTYFSSVFLGGVDSLVDLMNKSFPELGIKKTDCRQLSWIDTIIFYSGVVNY
_      .H.....TV.G....I.H.......................KEF.....T........F

     360       370       380       390       400       410
395641 DTDNFNKEILLDRSAGQNGAFKIKLDYVKKPIPESVFVQILEKLYEEDIGAGMYALYPYG
_      N.A..K..........KKT..S............TAM.K.........V.....V.....

     420       430       440       450       460       470
395641 GIMDEISESAIPFPHRAGILYELWYICSWEKQEDNEKHLNWIRNIYNFMTPYVSKNPRLA
_      ...E..............M.....TA...........I..V.SV...T.....Q.....

     480       490       500       510       520       530
395641 YLNYRDLDIGINDPKNPNNYTQARIWGEKYFGKNFDRLVKVKTLVDPNNFFRNEQSIPPL
_      ........L.KTNHAS..................N.......K................

     540
395641 PRHRH
_      .P.H.

Elapsed time:  0:00:00
```

GeneStream align Home Page

Query sequences and their profile-profile alignment

Download similarity matrix Download alignment

```
Query 1   >
          MNCSAFSFWFVCKIIFFFLSFHIQISIANPRENFLKCFSKHIPNNVANPKLVYTQHDQLYMSILNSTIQNLRFISDTTPKPLVIVTPSNNSHIQATILCS
          KKVGLQIRTRSGGHDAEGMSYISQVPFVVVDLRNMHSIKIDVHSQTAWVEAGATLGEVYYWINEKNENLSFPGGYCPTVGVGGHFSGGGYGALMRNYGLA
          ADNIIDAHLVNVDGKVLDRKSMGEDLFWAIRGGGGENFGIIAAWKIKLVAVPSKSTIFSVKKNMEIHGLVKLFNKWQNIAYKYDKDLVLMTHFITKNITD
          NHGKNKTTVHGYFSSIFHGGVDSLVDLMNKSFPELGIKKTDCKEFSWIDTTIFYSGVVNFNTANFKKEILLDRSAGKKTAFSIKLDYVKKPIPETAMVKI
          LEKLYEEDVGAGMYVLYPYGGIMEEISESAIPFPHRAGIMYELWYTASWEKQEDNEKHINWVRSVYNFTTPYVSQNPRLAYLNYRDLDLGKTNHASPNNY
          TQARIWGEKYFGKNFNRLVKVKTKVDPNNFFRNEQSIPPLPPHHH

Query 2   >
          MKCSTFSFWFVCKIIFFFFSFNIQTSIANPRENFLKCFSQYIPNNATNLKLVYTQNNPLYMSVLNSTIHNLRFTSDTTPKPLVIVTPSHVSHIQGTILCS
          KKVGLQIRTRSGGHDSEGMSYISQVPFVIVDLRNMRSIKIDVHSQTAWVEAGATLGEVYYWVNEKNENLSLAAGYCPTVCAGGHFGGGGYGPLMRNYGLA
          ADNIIDAHLVNVHGKVLDRKSMGEDLFWALRGGGAESFGIIVAWKIRLVAVPKSTMFSVKKIMEIHELVKLVNKWQNIAYKYDKDLLLMTHFITRNITDN
          QGKNKTAIHTYFSSVFLGGVDSLVDLMNKSFPELGIKKTDCRQLSWIDTIIFYSGVVNYDTDNFNKEILLDRSAGQNGAFKIKLDYVKKPIPESVFVQIL
          EKLYEEDIGAGMYALYPYGGIMDEISESAIPFPHRAGILYELWYICSWEKQEDNEKHLNWIRNIYNFMTPYVSKNPRLAYLNYRDLDIGINDPKNPNNYT
          QARIWGEKYFGKNFDRLVKVKTLVDPNNFFRNEQSIPPLPRHRH

Alignment
          1 MNCSAFSFWFVCKIIFFFLSFHIQISIANPRENFLKCFSKHIPNNVANPKLVYTQHDQLYMSILNSTIQNLRFISDTTPKPLVIVTPSNNSHIQATILCS  100
          1 MKCSTFSFWFVCKIIFFFFSFNIQTSIANPRENFLKCFSQYIPNNATNLKLVYTQNNPLYMSVLNSTIHNLRFTSDTTPKPLVIVTPSHVSHIQGTILCS  100

        101 KKVGLQIRTRSGGHDAEGMSYISQVPFVVVDLRNMHSIKIDVHSQTAWVEAGATLGEVYYWINEKNENLSFPGGYCPTVGVGGHFSGGGYGALMRNYGLA  200
        101 KKVGLQIRTRSGGHDSEGMSYISQVPFVIVDLRNMRSIKIDVHSQTAWVEAGATLGEVYYWVNEKNENLSLAAGYCPTVCAGGHFGGGGYGPLMRNYGLA  200

        201 ADNIIDAHLVNVDGKVLDRKSMGEDLFWAIRGGGGENFGIIAAWKIKLVAVPSKSTIFSVKKNMEIHGLVKLFNKWQNIAYKYDKDLVLMTHFITKNITD  300
        201 ADNIIDAHLVNVHGKVLDRKSMGEDLFWALRGGGAESFGIIVAWKIRLVAVPKSTMFSVKKI-MEIHELVKLVNKWQNIAYKYDKDLLLMTHFITRNITD  299

        301 NHGKNKTTVHGYFSSIFHGGVDSLVDLMNKSFPELGIKKTDCKEFSWIDTTIFYSGVVNFNTANFKKEILLDRSAGKKTAFSIKLDYVKKPIPETAMVKI  400
        300 NQGKNKTAIHTYFSSVFLGGVDSLVDLMNKSFPELGIKKTDCRQLSWIDTIIFYSGVVNYDTDNFNKEILLDRSAGQNGAFKIKLDYVKKPIPESVFVQI  399

        401 LEKLYEEDVGAGMYVLYPYGGIMEEISESAIPFPHRAGIMYELWYTASWEKQEDNEKHINWVRSVYNFTTPYVSQNPRLAYLNYRDLDLGKTNHASPNNY  500
        400 LEKLYEEDIGAGMYALYPYGGIMDEISESAIPFPHRAGILYELWYICSWEKQEDNEKHLNWIRNIYNFMTPYVSKNPRLAYLNYRDLDIGINDPKNPNNY  499

        501 TQARIWGEKYFGKNFNRLVKVKTKVDPNNFFRNEQSIPPLPPHHH   545
        500 TQARIWGEKYFGKNFDRLVKVKTLVDPNNFFRNEQSIPPLPRHRH   544
```

Score = -0.146E+03, Identities = 450/545(82%), Gaps = 1/545(0%)

FFAS03 scores below -9.5 usually mean significant similarity (less than 3% of false positives)
Conserved residues are shown in red

Dot plot representing regions of similarity of these two profiles

Current averaging radius is 1 (averaging window = averaging radius*2+1)
Change averaging radius to 1 (0-10) ✔ Show alignment [Redraw dot plot]

Download averaged similarity matrix Download postscript version